



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION IX
75 Hawthorne Street
San Francisco, CA

January 11, 2017

George ("Pat") Brooks
US Department of the Navy
33000 Nixie Way, Bldg 50
San Diego CA 92147

Dear Pat:

Thank you for the hard work and deep thought that went into crafting the Draft Radiological Data Evaluation Plan. I also appreciated the January 5, 2017, conference call and followup side conversations with you and Zach to ensure we understand relevant facts. We have shared these broad overall comments during those calls already, and we appreciate that the Navy has also already been considering these concerns and taking steps to address them. In addition to these broad comments, the attached more specific comments, will give deeper illustration of these categories of concerns.

1. The quality of prior data has varied due to differences in analytical methods and other reasons. Historical data should be screened carefully for reliability and to ensure that any comparisons are meaningful and consistent prior to spending significant time and resources on statistical analyses.
2. Where data quality is reliable, statistical tests are indeed another tool to find new evidence of potential falsification in previously undiscovered anomalies. However, prior data may also be unreliable because prior potential falsification has unknown scope. For example, earlier investigation and resampling has verified falsification of soil confirmation sampling. Other allegations have surfaced about potential disappearing of samples and tampering with data in spreadsheets. Because of unknown scope of prior falsification and the above challenges in data quality, statistical tests should not be used to automatically and definitively rule out locations to collect new samples in Phase 2. Sampling in many areas will be necessary regardless of the findings of the statistical tests.
3. Due to the above concerns, random confirmation sampling should be performed in addition to investigating outliers identified as a result of the Data Evaluation Plan and compared with previous results.
4. Phase 1 and Phase 2 should be iterative. Sampling does not have to wait until full statistical analysis. Findings from field sampling should help inform statistical analysis.

I look forward to continuing to work together on addressing this priority issue. Please contact me any time to discuss any of these comments. My phone number is 415-947-4187 and my email address is lee.lily@epa.gov.

Sincerely,



Lily Lee
Remedial Project Manager

Attachment

cc. Nina Bacey, State Department of Toxic Substances Control
Tina Low, Regional Water Quality Control Board
Amy Brownell, San Francisco Department of Public Health

ATTACHMENT –Draft Radiological Data Evaluation Plan, Former Hunter’s Point Naval Shipyard (HPNS), San Francisco, California, January 2017
USEPA Review dated January 11, 2017

These comments cover the data quality objectives (DQOs) as currently written. However further refinement of the DQOs will be required as the initial data analysis is being completed and as more information becomes available so that the efficiency and relevance of the proposed data evaluations/statistical analyses are maintained throughout the project.

1. The Draft Radiological Data Evaluation Plan, Former Hunter’s Point Naval Shipyard (HPNS), San Francisco, California (the Plan) does not propose conducting a data quality evaluation. The Phase I Data Evaluation proposes to organize data sets into groups for each survey unit/trench unit or work area, parcel, and survey type to run statistical evaluations. Once the data is grouped, it is recommended that a data quality evaluation be conducted prior to statistical data analysis to ensure data sets are sufficiently comparable (for comparative analyses) and are of sufficient quality for decision making. Such a data quality review may include, but is not limited to, the following:
 - a. Review of data sets to determine if a data validation was performed on the data
 - b. If data validation was not performed, ensuring each data set was generated under a defined quality assurance/quality control (QA/QC) program, and includes results of QC samples
 - c. Determining if detection limits (especially for Radium-226 [Ra-226] quantitated from gamma spectrometry analysis using the 186 kiloelectron volt [KeV] energy line) met the project requirements
 - d. For comparison evaluations, ensuring the data sets being compared were generated using the same analytical methods, data reduction methods, and detection limits.

This data quality review will ensure the statistical analyses are conducted on data sets with sufficient documentation regarding how the data was generated and are considered usable or comparable data so that the results of such statistical evaluations provide relevant and useful information. For efficiency, where data quality is not reliable or comparable, additional analysis may not be necessary or helpful.

2. Discussions held during the Navy Scoping conference call on January 5, 2017, indicated that much of the historical laboratory gamma spectrometry data from the on-site lab reported results for Ra-226 that was quantitated using the 186 KeV energy line, rather than using the ingrowth method and quantitating Ra-226 from the daughter radionuclide Lead-214 (Pb-214) at the 609 KeV energy line. Quantitation of Ra-226 using the 186 KeV energy line is subject to interference from Uranium 235 (U-235) gamma ray emissions that occur at approximately the same energy and may be skewed due to higher amounts of backscatter photons in this energy range. Ra-226 data generated using the 186 KeV energy line may produce false positive results and may or may not be useable

for decision making. Please clarify how Ra-226 data generated using the 186 KeV energy line will be handled.

3. The Phase 1 investigation should include some confirmation sampling of soil columns to test that the hypothesis of the statistical test used for Phase I for identifying outliers are valid. It would not be prudent to wait until after all of the Phase 1 and Phase 2 data reviews to begin confirmation sampling.
4. Page 1, Section 2 Data Evaluation Plan, Phase 1 Evaluation, Bullet 1: Please add release criteria for decisions, e.g. radionuclides and concentrations.
5. Page 1, Section 2 Data Evaluation Plan, Phase 1 Evaluation, Bullet 3
 - a. Sub-bullet 1: Statistical tests need to be better defined. Also, please state how the outliers will be calculated and presented. i.e. Will the data for each survey unit be presented with a “Box and Whisker Plot?”
 - b. Sub-bullet 2: We suggest the contractor provide example calculations for how this will be performed.
6. Page 2, Section 2 Data Evaluation Plan, Phase 2 Evaluation, bullet 6: A certain percentage of additional random samples should be collected when any anomalous outliers are investigated by confirmation sampling and compared with previous data sets.
7. Page 3, Section 3 Data Quality Objectives, Step 1 Identify the Problem, Identify Decision Making Method: Some additional field work should be performed for anomalous data and identified outliers for both Phase 1 and Phase 2.
8. Page 3, Section 3 Data Quality Objectives, Step 2 Identify the Decision, Bullet 1: Add release criteria decision radionuclides and concentrations. Based on release criteria alone, will outliers be identified that have an uncertainty that is a percentage of the release criteria for confirmation sampling?
9. Page 4, Section 3 Data Quality Objectives, Step 3, Phase 2:
 - a. Bullet 2: What percentage of archived samples will undergo a “Physical Review?” What happens if/when any archived samples do not have the same physical appearance of other samples in the survey unit or if some samples are missing?
 - b. Bullet 5: Random confirmation sampling should be performed in addition to investigating survey locations with outliers and compared with previous results.
10. Page 4, Section 3 Data Quality Objectives, Step 4: A list of prioritized survey units should be provided to the contractor in the order that the datasets should be investigated/assessed and confirmation sampling.

11. Page 4, Section 3 Data Quality Objectives, Step 5: The first paragraph should be revised to state that a data quality assessment will be required to ensure the data are usable for decision making in the data evaluation process. Further, it is not clear that it can be concluded that the data can be used to support decisions for transfer without confirmation sampling. Confirmation sampling should be done because these decisions should not be based solely on statistical evaluation. For example, it was stated during the January 5, 2017, conference call that it is not possible to identify whether only a few samples from an 18-sample set were falsified as the evaluation outlined in the Plan only can identify whether an entire 18-sample set were falsified. Because falsification could have been done to resolve apparent hot spots in a data set, it is possible that data for one to three samples in a data set were falsified. Since this cannot be identified using the proposed statistical analyses, confirmation sampling will be necessary. Please revise Step 5 to include a data quality assessment and confirmation sampling.
12. Page 4, Section 3 Data Quality Objectives, Step 5 Develop a Decision Rule, Phase 1 Decision Rules:
 - a. All data sets may not have a normal distribution (e.g. from individual survey units. For efficiency, in the case of non-normal distributions, additional analysis may not be helpful or necessary.
 - b. Summary Statistics, bullet 3: Likewise, radiological results reported as detections, but which include a counting or total propagated uncertainty (TPU) error that is greater than the reported result, should be considered un-reliable and flagged for investigation in Phase 2.
13. Page 5, Section 3 Data Quality Objectives, Step 5, Phase 1 Decision Rules, Bullet 2 Histograms
 - a. Sub-bullets 1 and 2: Please describe how outliers will be flagged.
 - b. Sub-bullet 4: We suggest striking “professional judgement” and adding how outliers will be mathematically identified.
14. Page 5, Section 3 Data Quality Objectives, Step 5, Phase 1 Decision Rules, Bullet 3 Normal probability plots:
 - a. Sub-bullet 1: This may not be possible since Ra-226 and eRa-226 are not identified separately. What is the definition for identifying outliers? 1st/3rd Quantile $\pm 1.5(IQR)$, 1st/3rd Quantile $\pm 3(IQR)$?
 - b. Sub-bullet 2: How will this be presented graphically? We suggest that the contractor present a “test case” from beginning of evaluation to the end and include decision points.
15. Page 5, Section 3 Data Quality Objectives, Step 5, Phase 1 Decision Rules, Bullet 4 Posting plots: How will outliers in the Final Status Survey results be defined? Will they be based on a survey unit or all parcel data? It is not clear that it will be possible to identify falsification from posting plots if a sample was moved to a location that was less likely to be contaminated within the triangular systematic survey grids.

16. Page 5, Section 3 Data Quality Objectives, Step 5, Phase 1 Decision Rules, Bullet 5 Time Series Plots, sub-bullet 1: Will a plot be generated for all radionuclides or just the radionuclides identified for clearance in the ROD? i.e. K-40 and other natural nuclide progeny.
17. Page 6, Section 3 Data Quality Objectives, Step 5, Phase 1 Decision Rules, Bullet 9 Two-Dimensional Paired Kolmogorov-Smirnov (Peacock) Test: The proposal to evaluate equilibrium comparisons should consider the following:
- a. The Two-Dimensional Paired Kolmogorov-Smirnov (Peacock) Test would be most effective when a reliable data set from Hunters Point Shipyard is used as the control data set, but this is not proposed. It is unclear how this test can be effective without a reliable control data set for comparison, particularly since there are questions about the reliability of the entire data set that is under evaluation. Please revise the Plan to include use of a reliable data set from Hunters Point Shipyard or explain how this test can be effective without one.
 - b. Offsite laboratory results may not provide a separation of Ra-226 activity concentration from eRa-226 concentration whenever an ingrowth method was used. The narrative of the data packages may be the only place where this could be found, but the narrative may not include which energy lines were averaged after ingrowth to quantify Ra-226 (Pb-214, Bi-214, or both). For the Ra-226 results generated from off-site laboratories, the data should be examined to determine if the sample was placed in a sealed container and stored for 21 days prior to analysis. If the 21-day ingrowth was not completed, an evaluation of how this was accounted for in the reporting of the final results should be investigated. Please ensure that the calculations and data sets used in such evaluations are presented to the team members to ensure all parties are in agreement with the approach. The Contractor will need to have copies of analysis libraries to see which energy lines are used for quantifying Ra-226. Ingrowth method may use Pb-214/Bi-214 lines; hence no Pb-214/Bi-214 results.
 - c. It is unclear why Th-232 results from site samples will be compared to Pb-212 and Bismuth 212 (Bi-212) results rather than using results from Actinium-228 (Ac-228), with a 6.1-hour half-life to evaluate the equilibrium status. Using the Pb-212 radionuclide to evaluate the equilibrium status would require a forty-year ingrowth time to establish secular equilibrium. If a ratio of Th-232 concentrations to Pb-212 concentrations is established to complete the comparison, information regarding the assumptions and correction factors applied will need to be provided. Please provide additional information regarding how this evaluation will be conducted. A comparison of Th-232 by alpha spec should be done with all daughter progeny by gamma spec (Ac-228, Pb-212, Bi-212 and Tl-208).
 - d. Please specify the statistical confidence level associated with a p value of 0.05.

18. Page 6, 3 Data Quality Objectives, Step 5 Develop a Decision Rule, Phase 1 Decision Rules, Bullet 9: For the Benford's Law Test, most first digits in the dataset would be expected to be 0 or 1 for many of the radionuclides, whether detected or falsified. For example, it is possible that a result of 3.4 picoCuries per gram (pCi/g) was modified to 1.4 pCi/g or even 0.4 pCi/g. It does not appear that this test or any of the statistical tests would detect such a change to an individual result. The second, third, and fourth integer tests may produce more meaningful results.
19. Page 6, 3 Data Quality Objectives, Step 5 Develop a Decision Rule, Phase 1 Decision Rules, Bullet 10 General Logic: How will sample locations within the survey unit be selected?
20. Page 6, 3 Data Quality Objectives, Step 5 Develop a Decision Rule, Phase 2 Decision Rules: During Phase 2, recounting selected archived samples that are determined to be consistent with the survey unit soil type should be done to validate the data set and conclusions from Phase 1. For example, it is unclear if there are cases where data, but not samples, were manipulated. Further, analyzing selected archived samples may help in evaluation of cases where only a few samples in an 18-sample data set were falsified. In addition, validated laboratory results with counting uncertainty/TPU and validation flags need to be evaluated. It is our understanding that the NIRIS database was not populated directly from laboratory EDDs. Where validated data is not available, then for efficiency, further analysis may not be helpful or necessary.
21. In the future, as the plans evolve, a decision tree should be provided so that it is clear how decisions will be made. For example, a decision tree would graphically illustrate when survey units/samples would be recommended for Phase 2.
22. In the future, as the plans evolve, please address buildings. More specifically, on December 7, 2016, EPA gave the Navy an updated version of its detailed technical recommendations. On page 7, Recommendation 7 stated "The Navy should also address the elevated levels found in Buildings 271 and 406. . . . The Navy should address the implications of these findings for these and other buildings." The calculations shown in the attachment raise questions about several areas of previous work conducted by Tetra Tech EC in buildings. Based on the information we have at this time, EPA would recommend some rescanning of buildings at the site. We therefore support the recommendations for rescanning from the State of California Department of Public Health (CDPH) letter in November, 2016. We understand that the Navy has chosen to evaluate buildings on a longer time frame and to focus on evaluating soil data first. When in the future more resources are available to study the buildings issue in more depth, we will look forward to revisiting this concern and reviewing any new information that may help us understand better any potential risks to public health and the environment related to buildings.